EXPLORING THE ROLE OF SPEECH IN THE TRANSMISSION OF AIRBORNE VIRUSES VIA A NOVEL METHOD FOR COMBINING AEROSOL AND PHONETIC ANALYSIS

CALEB EVERETT*[1], MARVA SEIFERT[2], CHANTAL DARQUENNE[2], RYAN TUMINELLO[3] and JONATHAN SLADE[3]

*Corresponding Author: caleb@miami.edu
[1]Department of Anthropology, University of Miami, Coral Gables, USA
[2]School of Medicine, University of California San Diego, La Jolla, USA
[3]Department of Chemistry and Biochemistry, University of California San Diego, La Jolla, USA

Aerosol particle emissions from speech are thought to play a key role in airborne pathogen transmission, though relatively little is known about the mechanisms through which articulatory gestures produce aerosol particles. We describe a novel method that allows for the more precise isolation of the production of aerosol particles during speech. Some of the particles are smaller than those detected with previous approaches. Our approach allows for the isolation of aerosol bursts that are associated with specific articulatory gestures. Combined aerosol and phonetic analysis illustrates how this approach can shed light on the mechanisms through which aerosols are generated during speech. We briefly discuss ongoing work that has substantially refined this method. In relationship to language evolution, this work highlights the possibility that certain articulatory gestures could have "maladaptive" characteristics, in that they might generate inordinate numbers of aerosols. We stress that more experimental research is required to address this possibility.

1. Introduction

The role of speech in the production of airborne particles and associated pathogen transmission has recently been highlighted as epidemiological evidence implicates speaking in SARS-CoV-2 transmission. (Meselson 2020, van Doremalen et al. 2020). Previous studies have utilized several methods to examine how speech produces such particles, including laser-based detection, airflow testing and aerosol analysis. (Stadnytskyi et al. 2020, Asadi et al. 2020) The latter is essential as growing consensus exists that particles smaller than 10 μm in diameter, often referred to as aerosols, can transmit a variety of pathogens. (Fennelly 2020) The SARS-CoV-2 virus appears to remain viable in aerosols, which can remain airborne for hours, further increasing the odds of transmission. (Fennelly 2020) Research with an aerodynamic particle sizer (APS) has allowed scholars to show that some speakers are "super-emitters" of aerosols and that some sound types correlate with higher emission rates. (Asadi et al. 2019, 2020) However, APS instrumentation only samples once per second, limiting the extent to which it can isolate the role of specific articulatory gestures in producing aerosol particles. Other methods that have been used face related limitations. In English, words typically last 150-500 ms, while the most common syllable types last about 120-260 ms and individual sounds 60-150 ms. (Greenberg et al. 2003) Given the durations of such units in English and other languages, a more complete understanding of aerosol generation via speech could benefit from an approach with higher temporal resolution in aerosol measurements.

Here we utilize a method for detecting aerosols, one not previously applied to speech in the manner we illustrate, to address two lacunae in our current understanding of particles generated during talking: i) the size distribution of aerosols from 0.07 μm in aerodynamic diameter (i.e., at a size resolution of single viruses) to 10 μm and ii) the mechanics of the production of such fine aerosol particles on timescales relevant to individual sound types and utterances. Addressing these lacunae is critical to elucidating the full range of emission variations across speakers and sound categories and could potentially help refine the modeling of speech-based pathogen transmission. We outline a new approach that allows us to isolate with relative precision the moments at which aerosols emerge from speakers' mouths and the concentrations of aerosols speakers produce at various size bands. Phonetic analysis in tandem with aerosol detection reveals relationships between aerosol emissions and certain sound types not accessible by previous approaches. In this work we detail initial results based on three speakers, given pandemic-related restrictions to the number of participants we could test. This general method is being applied to dozens of participants in the coming months, though in current work participants breathe in particle-free air, in contrast to the results described here. The results discussed here simply serve as an illustration of the overall method, demonstrating how the heightened temporal and physical resolution of the approach offers potential gains to our understanding of the aerosol emissions associated with speech.

2. Results and discussion

For this study, three participants read various stimuli and breathed (nasally) at a natural rate into an aluminum funnel (opening diameter of 20 cm and length of 25 cm), connected to an electronic particle impactor that could measure the size and volume of aerosols produced ten times per second. Each individual's nose and mouth were within the funnel opening. The funnel was attached to a Dekati electrical low-pressure impactor (ELPI) via flexible conductive tubing with an inner diameter of 1.2 cm and a length of 25 cm. The short residence time in the tube and use of conductive tubing led to no distinguishable particle losses. This method detected aerosols at 10 Hz across 14 size bands, between 0.006 μm and

10 μm, though we focus only on those bands greater than 0.07 μm since they are relevant to viral transmission. Experiments were conducted in the absence of synthetic particle-free air to mimic a real-world setting. Background aerosol concentrations in the room air were measured before and after speaking or breathing. The mean background was then subtracted from the data measured while speaking or breathing. Although this method controlled for the background room aerosol by subtraction, in ongoing work we are using an approach with particle-free air. (Participants breathe in particle-free air and their exhaled air is mixed with particle free-air as well, rather than room air.) The key difference between the background-subtraction results is that the latter yield a much greater number of measured particles. Nevertheless, the results discussed here illustrate some of the gains of the new method, particularly the high physical and temporal resolution of the approach. Further, some of the results discussed here are also evident in our ongoing work employing particle-free air. For instance, across both approaches whispered sounds produce a higher number of particles than sounds produced with normal voicing at low amplitude.

In addition to the aerosol analysis, we also analyzed the utterances of the speakers acoustically. An InnoGear Condenser Professional Cardioid Microphone was placed about five cm to the right side of the funnel. (In our ongoing work, this is also combined with airflow analysis.) Speakers were given instructions and stimuli displayed on a desktop monitor placed immediately behind the funnel, so that they could read without moving their heads. Recordings were made onto a notebook computer at a sampling rate of 44.1 kHz. Relevant wav files were analyzed with PRAAT. A few dozen words were recorded for all speakers, in addition to the same paragraph. For this paper, speakers were asked to read at a “normal” volume and cross-speaker amplitude was not found to vary substantially. We did not have speakers read at intentionally varied volumes, as in previous work that uncovered an association between increased amplitude and increased aerosol emissions. (Asadi et al. 2019) Across all three speakers we observed greater aerosol production during speech than during normal breathing. Likely, greater differences in the concentrations of aerosols emitted between speaking and breathing occur when speaking at a higher amplitude or when singing, and minor amplitude differences across speakers may contribute to some of the inter-speaker variability we observed. (Asadi et al. 2019)

With the high temporal resolution of the method, we were able to assess aerosol production from specific sound types as highlighted in Fig. 1 for one speaker. In Fig. 1A we observe that, after mask removal, there are aerosol bursts immediately following each named letter of the recited alphabet. In this case the bursts of aerosols (diameter-weighted, *D*, μm cm^{-3}) occur primarily across two size bands of roughly 1-2 μm. Figure 1B illustrates how the temporal resolution of the method allows us to match aerosol bursts with individual syllables and even specific sound types when words are produced several seconds apart. Previous research suggests certain sound types, most notably the [i] high-front vowel, are associated with increases in aerosols, as is increased amplitude. (Asadi et al. 2020) Here we focused on consonants, including two consonant types that have previously been suggested to be relevant to particle emission. Given the temporal resolution of the method, we were able to directly observe aerosol bursts following some consonants in isolated words, including word-final consonants that have not previously been investigated. As illustrated in Figure 1B, at least in some words there was an increase in aerosols immediately following affricate consonants such as the last sounds in "catch" and "*h*" and the first sound in "*g*". Figure 1C illustrates the variation in aerosol emissions across size bands during the articulation of one sentence previously used in research on the airflow produced during speech. (Abkarian et al. 2020) In this case there is an increase in aerosols shortly following the production of the [p^h], an aspirated bilabial plosive. This bilabial plosive has previously been shown to create an intense horizontal stream of air. (Abkarian et al. 2020) Some scholars have speculated that aspirated consonants like [p^h] may be relevant to airborne pathogen transmission via increased particle emission, but this is the first experimental evidence of any increase in aerosols associated with such aspiration. (Inouye 2003) We stress that it is very preliminary evidence, however, and must be replicated across many more participants with the new approach involving particle-free air. Initial results with the new approach are broadly consistent with those in Figure 1, suggesting that the association between aspiration and increased aerosols may be robust. Figure 1D further demonstrates the extreme concentration of aerosols that can be produced during some segments of speech for some speakers, in this case "shhh!", a lengthened voiceless postalveolar fricative.

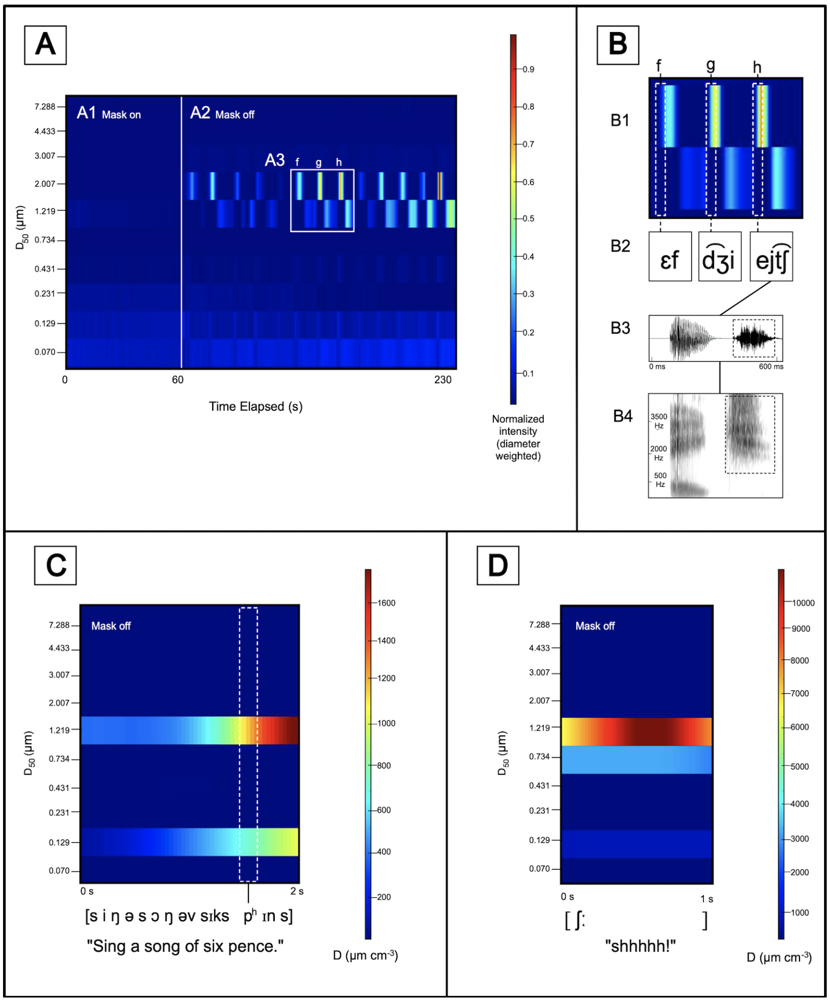

Figure 1. Temporal and physical resolution of method. Note that these values are much higher than the values obtained when speakers breathe in particle-free air, as demonstrated in our ongoing work. Panel A: Normalized heatmap of diameter-weighted aerosol particle concentration. A1. The speaker repeated the word "spar" at a normal volume, while wearing an N95 mask. A2. The speaker recited the alphabet at a normal volume, after removing his facemask. A3. Portion for *f*, *g*, and *h*, each said about eight seconds apart. Panel B: Section A3 enlarged, alongside phonetic details. B1. Aerosol heatmap. B2. Transcription of the three utterances with the International Phonetic Alphabet. B3. Waveform of [ejt͡ʃ]. Aperiodic sound waves associated with the voiceless affricate are highlighted. B4. Spectrogram of [ejt͡ʃ]. Aperiodic elements in high frequency range are highlighted. C. Aerosol visualization for one speaker's articulation of "Sing a song of six pence". D. Aerosols produced when the speaker was asked to say "shhh" (a voiceless postalveolar fricative), as though telling someone to be quiet. This was the most intense burst of aerosols for that speaker.

For more fluid segments of speech, the method still allows us to detect surges in aerosols associated with sounds or sequences of sounds though the aerosols do not occur in readily segmented bursts in such cases. It was observed that in fluid speech words with [st] sequences were weakly associated with increased emissions, an association that requires further exploration with a greater number of speakers and the particle-free approach. Such sequences, like the aforementioned affricates in "catch" and "shh" are of course voiceless. Previous work suggests that vocal cord vibration is also a key mechanism that dislodges and emits aerosol particles during speech (Asadi et al. 2019, 2020) but the high temporal resolution used here and in the laser-based detection in Stadnytskyi et al. (2020) suggests that at least some of the intense bursts of aerosols produced during speech are not produced in the larynx during vocal cord vibration. (In Stadnytskyi et al. (2020) it is observed that the voiceless interdental fricative in "healthy" emits many particles, though in a larger size range than that examined here.) It is possible that some sound types emit aerosols generated deeper in the respiratory tract, perhaps via the fluid-film burst mechanism in the terminal bronchioles. (Graham & Morawska 2009, Almstrand et al. 2010)

In short, the high resolution of this approach could be used to explore the detailed mechanisms through which aerosols are produced during speech, complementing other approaches. Ongoing refinement of this method could help to shed light not just on the temporal and physical dynamics of particle emission, but also on the mechanics through which aerosols are generated at the vocal cords and at other locations in the vocal tract during speech. Consistent with a growing literature using related methods that have more modest physical and temporal resolution, our results suggest that speaking does yield a high total number and volume of airborne particles that are potentially relevant to the transmission of some pathogens. Much work remains, however, to better understand how aerosols are produced during speech, along with the role of particularly articulatory gestures and associated sound types. In the context of language evolution, the preliminary results discussed here raise an interesting question: Do some articulatory gestures present a greater likelihood of intense aerosol bursts that can potentially transmit pathogens during an airborne pandemic? These results underscore this possibility, though we stress that much more experimental work is required, with a greater number of participants, to more fully understand whether some sound types have inaudible "maladaptive" features.

References

Abkarian, M., S. Mendez, N. Xue, F. Yang, H. Stone. (2020). Speech can produce jet-like transport relevant to asymptomatic spreading of virus. *Proc. Natl. Acad. Sci. U. S. A.* 117, 25237-25245

Almstrand, A., B. Baker, E. Ljungstrom, P. Larsson, A. Bredberg, E. Mirgorodskaya and A-C Olin. (2010). Effect of airway opening on production of exhaled particles. *J. Appl. Physiol*. 108, 584–8.

Asadi, S., A. Wexler, C. Cappa, S. Barreda, N. Bouvier, W. Ristenpart. (2019). Aerosol emission and superemission during human speech increase with voice loudness. *Sci. Rep.* 9, 2348.

Asadi, S., A. Wexler, C.D. Cappa, S. Barreda, N. Bouvier, W. Ristenpart. (2020). Effect of voicing and articulation manner on aerosol particle emission during human speech. *PloS One* 15, e0227699.

Fennelly, K. (2020). Particle sizes of infectious aerosols: Implications for infection control. *Lancet Respir. Med.* 8, 914-924.

Graham, R., L. Morawska. (2009). The mechanism of breath aerosol formation. *Journal of Aerosol Medicine and Pulmonary Drug Delivery* 22, 229-237.

Greenberg, S., H. Carvey, L. Hitchcock, S. Chang. (2003).Temporal properties of spontaneous speech – A syllable-centric perspective. *Journal of Phonetics* 31, 465-485 (2003).

Inouye, S. (2003). SARS transmission: Language and droplet production. *The Lancet* 362, 170.

Järvinen, A., M. Aitomaa, A. Rostedt, J.Keskinen, J. Yli-Ojanperä. (2014). Calibration of the new electrical low pressure impactor (ELPI+). *J. Aerosol Sci.* 69, 150-159.

Meselson, M. (2020). Droplets and aerosols in the transmission of SARS-CoV-2. *The New England Journal of Medicine* 382, 2063.

Stadnytskyi, V., C. Bax, A. Bax, P. Anfinrud. (2020). The airborne lifetime of small speech droplets and their potential importance in SARS-CoV-2 transmission. *Proc. Natl. Acad. Sci. U. S. A.* 117, 11875-11877.

van Doremalen, N., D. Morris, M. Holbrook, A. Gamble, B. Williamson, A. Tamin, J. Harcourt, N. Thornburg, S. Gerber, J. Lloyd-Smith, E. de Wit, V. Munster. (2020). Aerosol and surface stability of SARS-CoV-2 as compared with SARS-CoV-1. *The New England Journal of Medicine* 382, 1564-1567.